Homology-Independent Targeted Integration-mediated Gene Segment Replacement for Correction of *DMD* Mutations

Presented in partial fulfillment of the requirements for graduation “with research distinction in Neuroscience” in the undergraduate colleges of The Ohio State University

By Julian Havens

The Ohio State University April 2020

Project Advisor: Kevin Flanigan, Center for Gene Therapy at Nationwide Children’s Hospital

Acknowledgements

I would like to first thank all members of the Flanigan lab for the helpful discussions about this thesis and about life in general. The discussions had in the lab were instrumental in my growth as a scientist and as a human. Of particular note are two people: Tatyana Meyers, Ph.D. and Anthony Stephenson, Ph.D. Tatyana was very helpful throughout all microscopy imaging for which I have limited experience. Anthony has served as one of my primary mentors throughout much of the time I have spent in the lab and has provided me many opportunities to develop as a scientist. Without his guidance I would not have been able to achieve much of what I have accomplished thus far. Finally, I would like to thank Dr. Kevin Flanigan, who gave me the opportunity to conduct research in his lab and who has also been instrumental in my development as a scientist. Without the help and support of these and others, I truly would not be where I am today, and for that I am very grateful.

Abstract:

Duchenne muscular dystrophy (DMD) is an X-linked genetic disorder which causes muscle degeneration leading to loss of ambulation, cardiac dysfunction, respiratory complications, and premature death. The *DMD* gene is the largest gene within the human genome, and thus has a diverse mutational profile which prevents many corrective therapies from reaching a large patient cohort. My goal was to determine the feasibility of a new CRISPR/Cas9-based gene editing technique known as Homology Independent Targeted Integration (HITI) in replacing segments of the *DMD* gene. By replacing large segments of the *DMD* gene, a larger patient cohort could potentially be encompassed by a single therapy. As a proof of principle, I tested whether small (~1 kb) and medium sized (~175 kb) replacements were feasible at the 5' end of the *DMD* gene using a HITI-mediated replacement with previously validated Cas9 guide RNAs (gRNAs). Genomic DNA PCR was used to detect the expected knock-in while sequencing confirmed that the integration was seamless with no insertion or deletions (indels). I then optimized this system by altering the molecular ratios of the two HITI gene editing components (CRISPR/Cas9 and HITI donor DNA plasmid) and found that a 1:1 ratio was optimal. Based on these studies, small and medium HITI replacements are feasible within the *DMD* gene, thus a large HITI replacement (~715 kb) of *DMD* exons 41-55 (which would potentially correct ~37% of patients) was designed. This consisted of two plasmids encoding *i*) Cas9 with a GFP reporter and *ii*) a HITI donor DNA with the coding sequence (CDS) of *DMD* exons 41-55, two gRNAs, and an RFP reporter. Knock in of the CDS of *DMD* exons 41-55 was detected and analyzed by genomic DNA PCR and Sanger sequencing. As with the proof-of-concept studies, the *DMD* exon 41-55 CDS was seamlessly integrated in place of the natural *DMD* exon 41-55 locus. These studies lay the groundwork for development of a HITI-mediated gene correction therapy that could potentially restore full-length dystrophin in a large patient cohort.

Table of Contents:

Chapter 1: Background

1.1 Dystrophin and Duchenne Muscular Dystrophy

Duchenne muscular dystrophy (DMD) is an X-linked genetic disorder caused by myriad mutations within the *DMD* gene which contains a total of 79 exons and codes for the 427 kDa muscle isoform of the dystrophin protein (Fig. 1).[1] Dystrophin is a structural protein which serves to reinforce the plasma membrane via a connection between cytoskeletal actin filaments and the dystroglycan complex (DGC) (Fig. 1A).[2] As such, dystrophin has several key domains including an N-terminal actin binding domain, a central rod domain comprised of spectrin-like

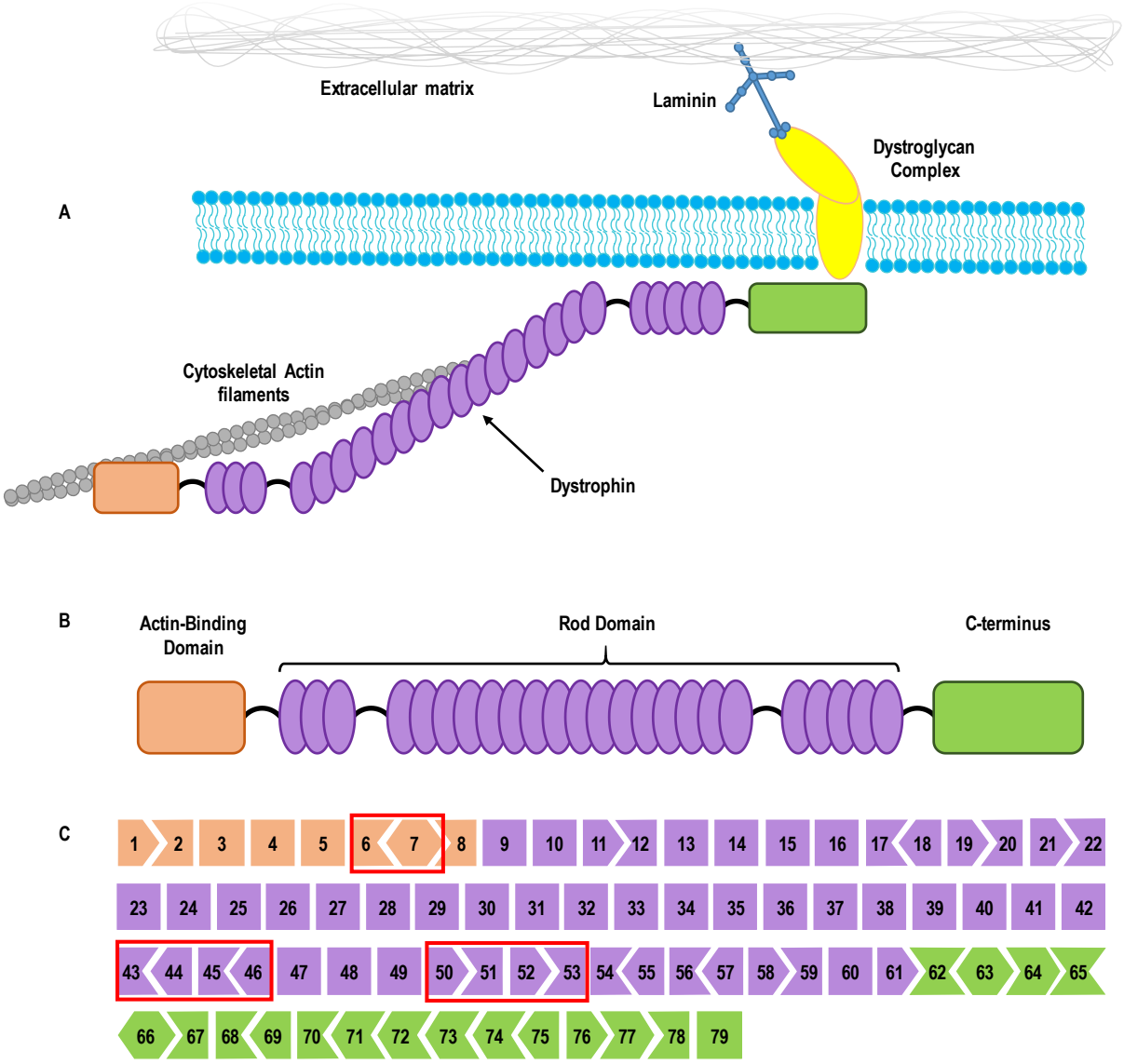

Fig. 1 (**A**) Schematic of the axis of force transduction in muscle cells. Dystrophin links the cytoskeletal actin to the transmembrane dystroglycan complex thus linking the cytoskeleton to the extracellular matrix via laminin. (**B**) Schematic of the dystrophin protein with the major domains labeled. (**C**) *DMD* gene diagram of exons corresponding to each domain in dystrophin. The shape of each exon depicts reading frame phasing, while red boxes show mutation hotspots within the *DMD* gene.

repeats with a second actin binding domain, and a C-terminal domain that directly interacts with the DGC (Fig. 1B).[2] Dystrophin acts as a shock-absorber during normal muscle contraction and is required to prevent muscle damage and degeneration during normal activity.[2] In the absence of dystrophin, muscle degeneration leads to weakness which eventually progresses to a loss of ambulation in the early teens.[1] Once in a wheelchair, patients have steep declines in cardiac and respiratory function (due to the involvement of the heart and diaphragm, respectively) which are the primary causes of the early mortality characteristic of DMD.[1]

The *DMD* gene has a diverse mutational profile, due in part to the size of the gene.[3] Single Nucleotide point mutations, which is the result of single base pair changes in the DNA sequence, account for about 10.5% of DMD causing mutations.[3] Exonic duplications account for about 10.9% of DMD mutations and occur when a portion of the gene is duplicated and placed directly adjacent to the original gene fragement.[3] Exonic deletions are when a portion of the gene containing one or more exons is fully excised from the gene, and account for about 68.5% of DMD mutations.[3] Both exonic deletions and duplications usually result in frameshift mutations that generally lead to loss of functional dystrophin protein. Another 6.9% of DMD mutations consist of subexonic insertions and deletions (indels) that also generally result in frameshift mutations.[3] Another 2.7% of DMD mutations consist of mutations that affect the splice sites of certain exons.[3] The final 0.5% of mutations consist of variable and highly specific mutations throughout the intronic regions of the DMD gene.[3] Despite this extensive mutational profile, gene editing has shown great potential in correcting many of the types of mutations described above.

1.2 CRISPR/Cas9 Gene Editing

Clustered Regularly Interspaced Short Palindromic Repeats and the associated protein 9 (CRISPR/Cas9) is an adaptive immune system found in bacteria that utilizes an RNA-programmable endonuclease to protect bacteria against viral invaders. This system, which consists of a guide RNA (gRNA) and a Cas9 endonuclease protein, has been repurposed to make precise double stranded breaks (DSBs) at a site complementary to the gRNA and near a short recognition sequence known as a protospacer adjacent motif (PAM) site (Fig. 2).

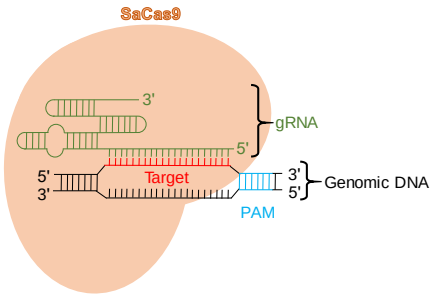

Fig. 2 Depiction of Cas9 targeting via the use of a gRNA complementary to a portion of genomic DNA. The genomic DNA bound by the gRNA is shown in red. In blue is the PAM site. The SaCas9 protein is shown in orange.

There are several different homologs of the Cas9 protein from different bacteria which have differences in size and PAM recognition sequence. The most well characterized variant is Cas9 from *streptococcus pyogenes* (SpCas9) which is encoded by 1,371 amino acids and has a PAM recognition sequence of 5'-NGG-3'.[4-6] A less commonly used Cas protein is from *staphylococcus aureus* (SaCas9) which, in contrast to SpCas9, is encoded by 1,053 amino acids and has a PAM recognition sequence of 5'-NNGRRT-3'.[7] The use of the smaller SaCas9 protein is preferable in virally delivered gene therapies on account of the limited cargo space (~5 kb) associated with viral vectors such as the Adeno-Associated Virus (AAV).[8]

1.3 Current and Future Clinical Approaches to DMD

1.3.1 Current Approaches

There is currently no cure for DMD. The contemporary treatment paradigm states that patients with DMD should be started on corticosteroids as soon as possible, which has been shown to increase the duration of ambulation by 2 to 4 years.[9] There is also some evidence that the heart benefits from early use of corticosteroids.[10] Current recommendations indicate that an ACE inhibitor or an angiotensin receptor blocker (ARB) should be started before the age of ten to prevent cardiac degeneration and subsequent dysfunction in patients with DMD.[11] Patients are eventually bound to a wheelchair as they lose their ability to walk, and due to the involvement of the diaphragm, patients will experience respiratory decline and will eventually require the use of a ventilator.[1]

1.3.2 Gene replacement therapies

One commonly used methodology for the correction of genetic disease in which there is a lack of functional protein is to deliver the coding sequence (CDS) of the protein with a tissue-specific promotor to ensure it is expressed only in the appropriate tissues.[12-14] AAV is the most common delivery system used for gene replacement in humans, but it's packaging limit is small (~5 kb), thus preventing delivery of the complete dystrophin CDS of ~11 kb. Instead, delivery of synthetic, miniaturized isoforms of dystrophin which lack non-essential domains has been shown to improve symptoms in DMD animal models and is currently being tested in human trials.[14,15] Despite this promising approach, the long-term benefits of miniaturized dystrophin isoforms are unknown and a therapy to restore full-length dystrophin would be preferred.

1.3.3 Exon skipping and reframing

Another method for correcting certain DMD-causing mutations is to induce exon skipping during pre-mRNA maturation with antisense oligonucleotides to restore a coherent reading frame. Two phosphorodiamidate morpholino oligomer (PMO) drugs have been approved that utilize this method which are eteplirsen and golodirsen which induce skipping of exons 51 and 53, respectively, potentially benefiting a combined 22% of patients.[16,17] The restoration of dystrophin and subsequent therapeutic benefit of PMOs are exceedingly low and would require weekly treatments to maintain therapeutic benefits.[18] Thus, there is a need for more robust therapies.

Recently, it has been shown that gene editing can be used to reframe mutant *DMD* genes with higher efficiency and potentially more long-term benefits. The most notable example is with exclusion of exon 51 which was demonstrated in a DMD dog model.[19] This approach targets the 3' splice elements to introduce inactivating mutations that result in exon 51 exclusion during pre-mRNA maturation. This results in a truncated, mutant form of dystrophin for which the long term benefits have not been adequately characterized, once again showing the need for a therapy that would restore full-length dystrophin.[19] This approach also suffers from being highly specific towards one type of DMD mutation that only has the potential to help a specific subset of the patient population (~13%), therefore a novel therapy that would be capable of correcting a wider range of mutations is needed.

1.3.4 Efforts to restore full-length dystrophin

Homology Directed repair (HDR) is the classically used method for integration of a desired DNA sequence into a genome and has been used to attempt restoration of the full-length dystrophin CDS, and thus full-length protein.[20,21] This methodology works by providing a donor DNA molecule encoding a knock-in sequence flanked by long regions of sequence homologous to the genomic context on either side of a DNA break induced by a nuclease (*e.g.* Cas9). The large homologous sequences drive the HDR DNA repair pathway over other DNA repair pathways in many proliferative cell types.[22] One study used this to induce knock-in of exon 44 in mutated induced pluripotent stem cells (iPSCs) that were derived from DMD patients. Although this correction was highly efficient, iPSCs are expected to poorly recapitulate the cellular environment of mature, non-proliferative muscle cells.[20] Another study attempted to correct a nonsense mutation within exon 53 in a mouse model by means of HDR using a dual-gRNA CRISPR/Cas9 approach and a donor containing the corrected exon 53 via an intramuscular injection into the tibialis anterior muscle.[23] Despite there being some evidence of

successful HDR, there was exceedingly low efficiency (~0.18% of total genomes) depicting a lack of potential for HDR to be used for an *in vivo* gene therapy and necessitating the design of a novel and more efficient method of knock-in for non-proliferating cells.[23]

1.4. Homology Independent Targeted Integration

Though HDR works well for integration in some cells, it has notoriously low efficiency in non-proliferating cell types such as muscle fibers, where the non-homologous end-joining (NHEJ) DNA repair pathway predominates.[21,24] Recently, a method was described which accomplishes high efficiency knock in using the NHEJ pathway called Homology-Independent Targeted-Integration (HITI).[25] HITI requires two components; *i*) CRISPR/Cas9 and *ii*) a donor DNA containing CRISPR/Cas9 cut sites flanking the desired knock-in fragment (Fig.3).[25] CRISPR/Cas9 generates a genomic DNA break while also cleaving the donor DNA and activating it as a NHEJ substrate for integration into the genome DSB. HITI was initially designed with a single target site and has been utilized to knock in a missing exon in *Mertk*, the gene implicated in a rat model of retinitis pigmentosa.[25] The HITI treated rats showed improved eye function when compared to their untreated or HDR treated counterparts.[25] This same study used mice to show that the *in vivo* efficiency of systemic AAV-mediated HITI knock-in of a reporter gene to be ~10% in the quadriceps muscle and ~3-4% in heart muscle.[25] This HITI approach shows the potential to be therapeutically beneficial as a gene therapy, and despite its apparent low efficiency, shows much more promise than the previously used HDR method of knock-in gene correction, especially in non-proliferating cells.[25] HITI also remains un-optimized, allowing for further improvements. Thus, this strategy could theoretically be used for the replacement of aberrant exons within the *DMD* gene.

Chapter 2: Feasibility Studies for HITI Exon Replacement

2.1 Introduction

The previously described HITI methodology has been utilized not only for the insertion of missing exons and reporter genes at a single site, but also for the replacement of small (~1.3 kb) portions of the *CCAT1* gene in human cancer cells.[26] To ensure that a similar approach would work at the *DMD* locus, previously validated gRNAs targeting up- and down-stream of exon 2 were utilized to remove this exon and subsequently knock in an exogenous DNA sequence (Fig. 3A). I also sought to determine whether larger replacements were feasible using HITI. To this end, a set of previously validated gRNAs, one upstream of exon 2 and one downstream of exon 3 were utilized in a deletion and subsequent HITI experiment to confirm the feasibility of larger genomic replacements than previously described (Fig. 3B).

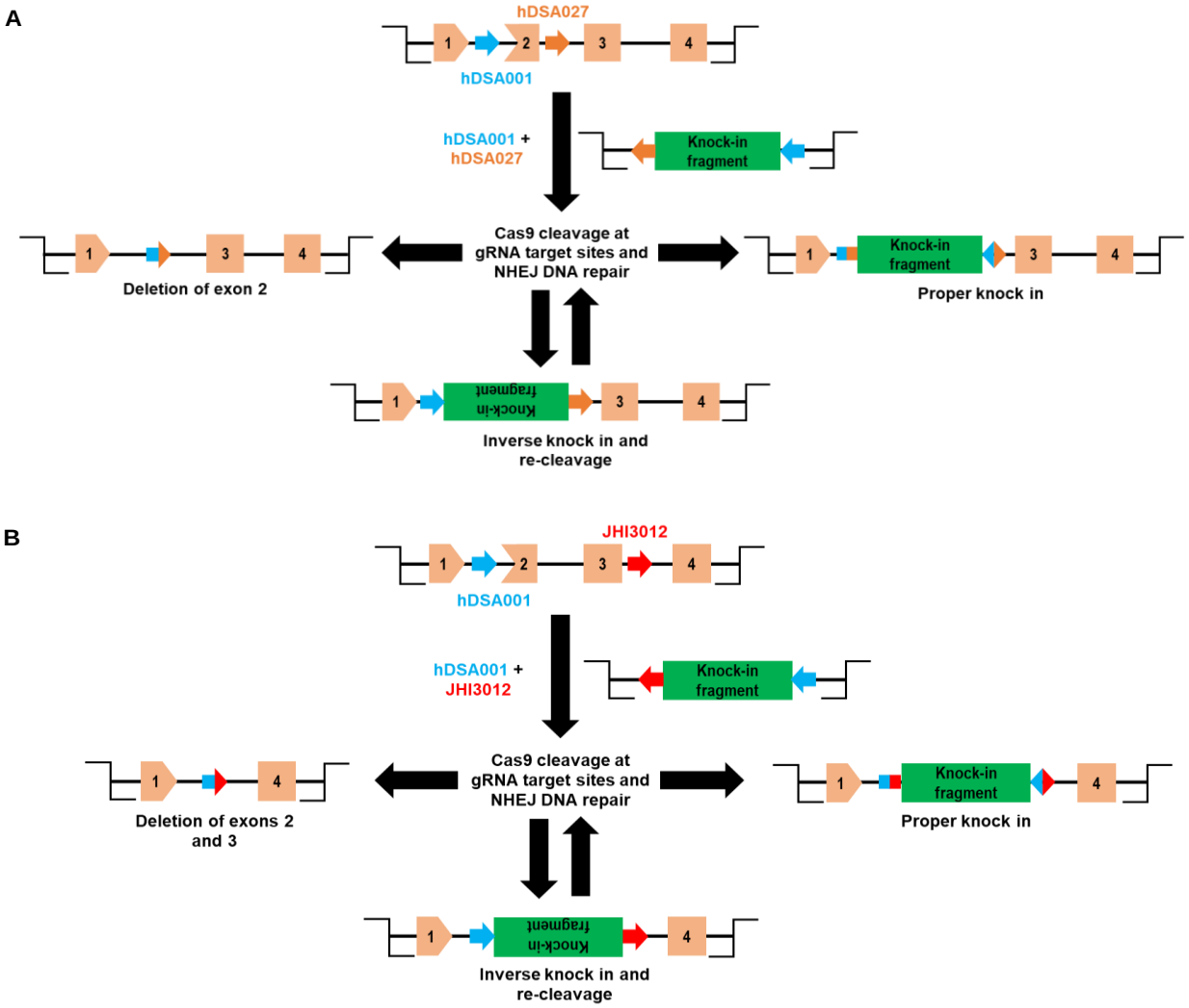

Fig. 3 Schematic showing the HITI strategy for (**A**) exon 2 replacement and (**B**) exon 2 + 3 replacement. Proper cleavage at the two genomic loci and of the knock in fragment can result in one of three possible outcomes. Simple deletion of the flanked exon(s), inverse integration of the knock in fragment, or proper forward knock in. Inverse knock in results in reconstitution of the cut sites allowing for re-cleavage.

For the replacement of exon 2 (small replacement, ~1 kb), two gRNAs flanking this exon were utilized to cut within the genome as well as the HITI donor vector. The cut sites on the donor vector were engineered to be the reverse complement of those cut sites in the genomic context and placed at opposite 5' and 3' ends of one another (Fig. 3A). This was done so that in the case of inverse integration of the HITI donor fragment, the Cas9 cut sites would be reconstituted, allowing for re-cleavage and a greater proportion of integrations being in the forward orientation (Fig. 3A). A similar strategy was used for the replacement of exons 2 and 3 (medium replacement, ~175 kb) with a gRNA targeting upstream of exon 2 and one downstream of exon 3 as well as a similarly designed HITI donor fragment (Fig. 3B).

These experiments utilized a “triple plasmid system” wherein one plasmid had encoded the exogenous knock-in donor DNA sequence flanked by two gRNA cut sites, and two additional plasmids encoded SaCas9 and the two gRNAs utilized in cutting the genomic DNA and the donor plasmid.

2.2 Materials and Methods

2.2.1 Molecular Cloning

Generation of the plasmids used in these studies was accomplished through several different traditional and modern cloning techniques. For the swapping of gRNAs, a technique known as restriction free cloning (RFC), which utilizes PCR to amplify the entire plasmid with two mega-primers that contain the desired change flanked by two regions of complementarity to DNA context surrounding the change, was used. All other cloning was accomplished with the In-Fusion cloning kit (Takara Bio) according to the manufacturer’s recommendations.

2.2.2 Cell culture and treatments

Human embryonic kidney 293 (HEK293) cells were cultured in HEK complete medium (Dulbecco’s modified Eagle medium high glucose supplemented with 10% cosmic calf serum, 1% 100X antifungal/antimicrobial and 1% 100X modified Eagle medium non-essential amino acids) in 10 cm^2 dishes until they were ~80-90% confluent. They were then dissociated from the dishes using 0.025% trypsin-EDTA and counted with a hemacytometer. Cells were plated in each well of a 12-well dish (200,000 cells/well) and allowed to grow for one day before being used for experimentation. Cells were cultured at 37°C with 100% humidity and 5% CO_2.

Transfection of cells with plasmid DNA was accomplish with the use of Lipofectamine LTX per the manufacturer’s recommendations with the indicated amount of DNA, 5 µL of Lipofectamine LTX, and 1 µL of Plus Reagent per µg DNA. Once transfected, cells incubated for

6 hours before the medium was replaced with fresh HEK complete medium. Cells were cultured for three days before extraction of genomic DNA for analysis.

2.2.3 Genomic DNA PCR analysis

Polymerase Chain Reaction (PCR) was performed using Q5 high-fidelity 2X master mix per the manufacturer's recommendations with the indicated primers and addition of 0.08 U/µL of Q5 Hot Start High-Fidelity DNA Polymerase. PCR products were resolved using electrophoresis on either 1% agarose-TAE or 10% polyacrylamide-TBE gels as indicated and stained with ethidium bromide. Images were collected using a BioRad ChemiDoc Imaging System with automatic optimal exposure times. The primers utilized for the PCR differed by assay. In the assay performed in chapter *2.3.2* the 5' end amplicon was generated with a forward primer that anneals upstream of the 5' Cas9 cut site and a reverse primer that anneals within the knock-in fragment. The 3' end amplicon was generated with a forward primer that anneals within the knock-in fragment and a reverse primer that anneals downstream of the 3' Cas9 cut site. Finally, in section *2.3.3*, the bulk amplicon was generated by the 5' forward primer and the 3' reverse primer from the previous assay, which encompasses the entire knock-in. Primer Tm's were calculated based on NEB's Q5 DNA polymerase 2X master-mix online suggestions and the extension time was calculated based on the amplicon length, utilizing 30 seconds per kilobase of amplicon.

2.2.4 EnGen Mutation Detection Assay

The EnGen mutation detection assay from NEB used for the confirmation of active gRNAs makes use of the T7 endonuclease I (T7E1) enzyme to cleave at mismatched DNA. To begin, genomic DNA PCR was used to generate amplicons surrounding the expected site of editing. These amplicons were gel purified and subsequently incubated at 95°C followed by slow annealing at a ramp speed of -0.1°C/second to allow for the reannealing of heterogeneous DNA indicative of editing. Next, the T7E1 enzyme was added to the reannealed DNA and allowed to incubate at 37°C for 15 minutes. The resulting cleaved DNA was analyzed via polyacrylamide gel electrophoresis (PAGE) stained with ethidium bromide.

2.3 Results and Discussion

2.3.1 Exon 2 and 3 targeting gRNAs

Prior to the experiments performed in this study, gRNAs targeting exons 2 and 3 were designed *de novo* by first identifying SaCas9 PAM sequences (5'-NNGRRT-3') within 1000 base pairs (bp) of the targeted exon as deletion and duplication mutations that affect a given exon

have a higher probability of also including the surrounding intronic sequence that is closest to the exon. Importantly, intronic targeting is preferred because the indels that are common with the NHEJ DNA repair pathway are less likely to be deleterious in non-coding regions. It was noted in the design of exon 2 targeting gRNAs that upstream targeting sequences tended to have much larger off-target profiles likely due to homogeneity of 3' splice elements, thus the gRNAs were designed both upstream and downstream of exon 2 while they were only designed downstream of exon 3.

Exclusion criteria were used to ensure optimal candidate gRNAs. Firstly, gRNA sequences containing putative RNA polymerase III termination signals (four or more contiguous thymidine residues in the coding strand) were excluded as this could lead to pre-mature termination during transcription from the U6 promoter. Next, gRNAs with more than 30 predicted off-target sites or any number of exonic off-target sites in the human genome as predicted by CCTop bioinformatics software were eliminated.[27] The predicted off-targets of the remaining gRNAs were noted and this information will be utilized to aid in analyzing off-target profiles at a later time. Finally, as mismatches between the target DNA and gRNA or suboptimal PAM sequences can inhibit gene editing, gRNAs were rejected if their target sequence or PAM contained single nucleotide polymorphisms (SNP) or variations (Vars) with greater than one percent minor allele frequency based on the ClinVar and dbSNP databases.[28,29]

After the initial screening, the candidate gRNAs were cloned into plasmids downstream of a U6 promoter along with an SaCas9 expression cassette driven by the cytomegalovirus immediate early enhancer and promoter (CMVP) for high-level, constitutive expression. These plasmids were used to transfect HEK293 cells to test for efficient cleavage at the expected genomic loci. Amplicons were generated from the genomic DNA flanking the sites of expected editing and the EnGen mutation detection kit, which makes use of the T7 endonuclease I (T7EI) enzyme that cleaves at sites of DNA mismatches, was used to check for proper editing. Through this experimentation it was revealed that the lead gRNA candidates for upstream exon two targeting was hDSA001, for downstream exon two targeting was hDSA027 and for downstream exon three targeting was JHI3012.

2.3.2 HITI Replacement of small and medium sized DMD *gene fragments*

To begin, a DNA fragment already housed within the lab was used to create a HITI donor vector which contained the genomic Cas9 cut sites on either end of the knock-in fragment as described above (Fig. 3). To test whether or not small and medium sized fragments of the *DMD* gene could be replaced with a HITI donor, HEK293 cells were co-transfected with three

plasmids: two containing the SaCas9 expression cassette and the chosen gRNAs, and a HITI donor plasmid containing paired gRNA cut sites flanking a GFP expression cassette. The gRNA pair for the small replacement was hDSA001 and hDSA027 while the gRNA pair for the medium replacement was hDSA001 and JHI3012 (Fig. 3).

Gel images from genomic DNA PCR for the exon 2 replacement (small replacement of ~1 kb) revealed that proper integration did occur at the expected locus when using knock-in specific primers (Fig. 4A). A similar result was noted for the replacement of exons 2 and 3 (medium replacement of ~175 kb) using a similar method for PCR (Fig. 4B). For both experiments, the expected knock-in bands were not present in cells treated with CRISPR only or with donor only showing that these components alone are not sufficient for knock in. The cells treated with a combination of CRISPR and a non-template donor (an identical donor lacking the Cas9 cut sites) also did not show the expected knock-in bands, confirming that Cas9 cleavage of the donor is necessary for HITI knock in to occur. The expected band from the exon 2 and 3 replacement was extracted and sequenced, confirming a seamless integration at the 5' and 3' ends of the insertion (Fig. 4C and 4D).

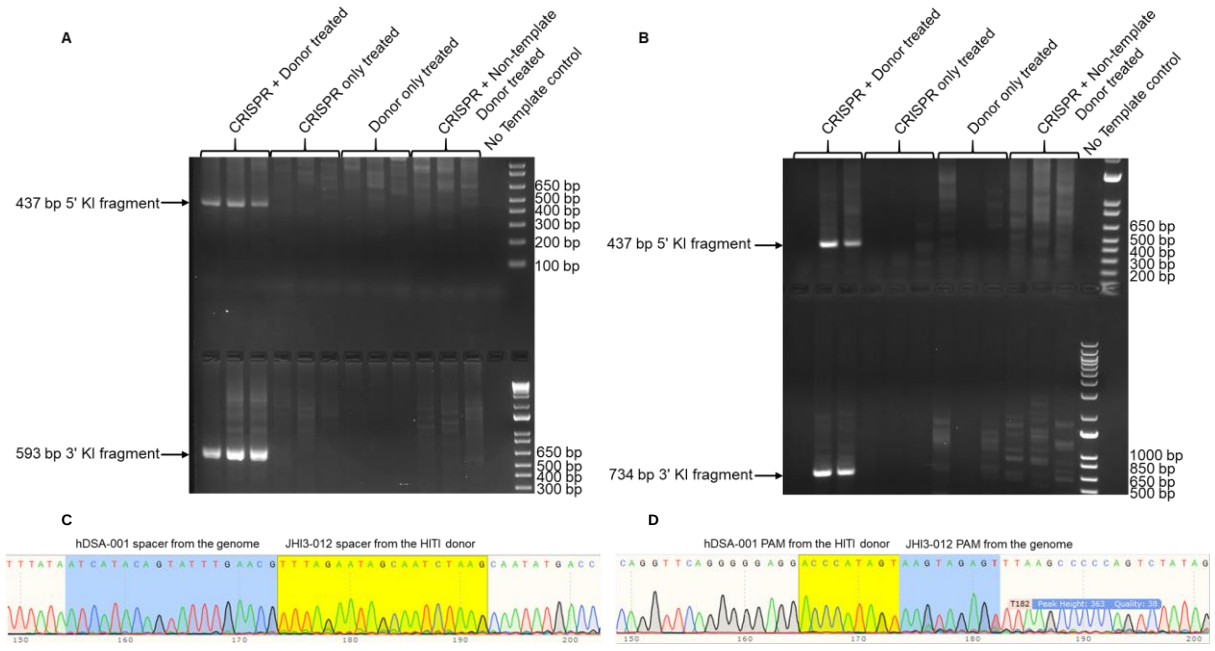

Fig. 4 Gel images showing successful knock in of the HITI donor following PCR of treated HEK293 genomic DNA with knock-in specific primers. This was shown for both (**A**) exon 2 replacement and (**B**) exon 2-3 replacement. Sequencing data showed that there was perfect joining of (**C**) the 5' end of the HITI insertion as well as (**D**) the 3' end of the HITI insertion.

2.3.3 Optimization of HITI plasmid ratios for small and medium sized replacements

Next, I wanted to ensure that the optimal ratios of Cas9 plasmid to Donor plasmid were being used to get the most deletion and integration events possible. To this end, variable amounts of Cas9:Donor plasmid were used to transfect HEK293 cells and subsequently screened using the PCR conditions described in section 2.3.2 (Fig. 5A). The experiment with decreasing Cas9:Donor ratios was completed with the small replacement and showed that the most integration occurred when there was a 1:1 ratio (Fig. 5A). As this was expected to be applicable regardless of the size of replacement, it was decided that the experiment with increasing Cas9:Donor would be conducted with the medium sized replacement (Fig. 5B). This PCR was also conducted with primers flanking the whole knock-in region instead of the knock-in specific primers to test whether detection of the entire knock-in locus was possible (Fig. 5B). The gel images indicated that the optimal ratio was 1:1 and showed that detection of the whole knock-in amplicon was possible, albeit at a lower efficiency than the knock-in specific primers (Fig. 5B). I hypothesize this lack of efficiency is due to amplification of the smaller deletion amplicon with higher efficiency than the full knock-in amplicon. It was believed that the limiting factor in this experimental set-up is the co-delivery of three plasmids, therefore, a dual-plasmid system was used for subsequent development of this potential therapeutic strategy.

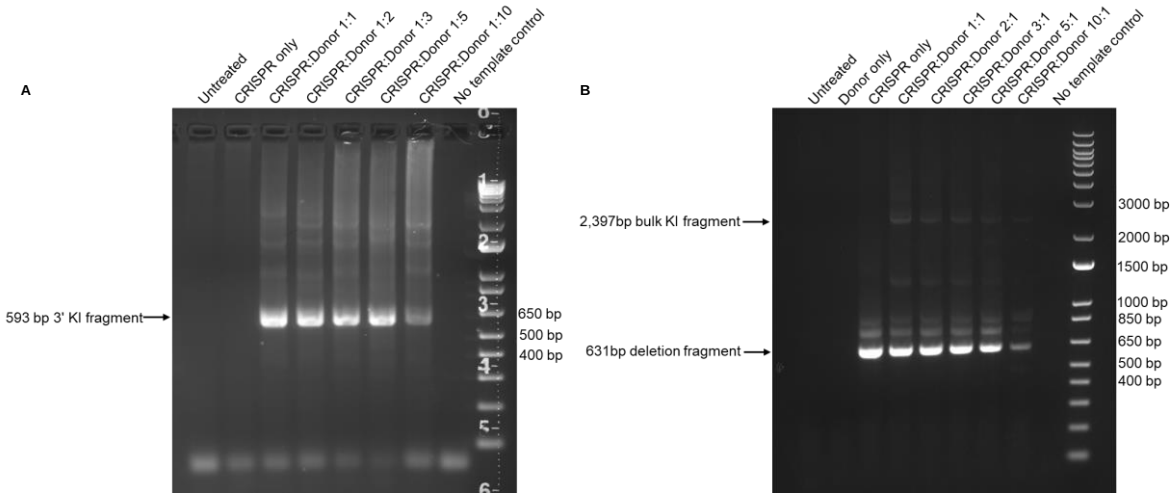

Fig. 5 Gel images depicting the CRISPR:Donor titration experiments. Results from both (**A**) increasing the Donor amount compared to the CRISPR amount and (**B**) increasing the CRISPR amount compared to the Donor amount. These experiments indicated a ratio of 1:1 is optimal for this triple plasmid co-transfected system.

2.3.4 Conclusions

These preliminary experiments depict that the HITI replacement of gene fragments is feasible within the *DMD* gene, both on a small (~1 kb) scale and on a larger (~175 kb) scale via the utilization of SaCas9, two gRNAs and a HITI donor fragment that contains the genomic

Cas9 cut sites on either end and in the orientation described in section **2.1**. This answer allowed for the optimization of the ratios of plasmid at 1:1 of Cas9 plasmids to donor plasmid, which could prove useful in cases that necessitate the use of a triple plasmid delivery or for the future of the dual-plasmid transfection that will be discussed in **Chapter 3**.

Chapter 3: HITI Replacement of *DMD* Exons 41-55

3.1 Introduction

Using the basic methodology described in **chapter 2**, I sought to design a large (~715 kb) HITI-based gene editing strategy that would enable the restoration of full-length dystrophin in a greater number of patients. To this end, bioinformatics analysis was done on the *DMD* gene to pick an efficient target for exonic replacement. Exons 41-55 encompass two mutational hotspots and efficient replacement of this region could benefit ~37% of DMD patients, and thus this was the target chosen for this study (Fig. 1C).[30] In the case that deletion, but not integration, occurs within the region, an open reading frame would be maintained creating a truncated, potentially therapeutic isoform of dystrophin analogous to the synthetic, miniaturized isoforms of dystrophin discussed in section *1.3.2*. Though the excision of 41-55 would truncate two different spectrin-like repeats, SWISS-MODEL online homology-modelling server was used to predict the structure of the resulting hybrid spectrin-like repeat that is formed from the joining of exons 40 and 56. The predicted hybrid spectrin-like repeat modelled on a helical bundle structure similar to the endogenous spectrin-like repeat 22 based on global quality estimates (Fig. 6).[2,31-35]

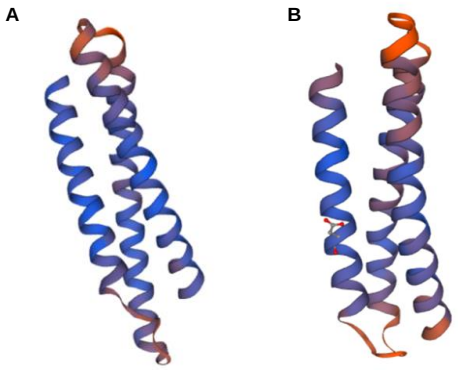

Fig. 6 Homology models of predicted structure of (**A**) the endogenous spectrin-like repeat 22 and (**B**) the hybrid spectrin-like repeat produced from joining of exons 40 and 56 built using SWISS-MODEL. Blue areas depict more favorable global quality estimates whereas red areas depict less favorable global quality estimates.

The HITI donor vector was designed such that the CDS of exons 41-55 (~2.5 kb) was placed between two regions of ~100 bp of endogenous intronic sequence to include the 5' and 3' splice elements. Just past the intronic sequence on both sides were placed the genomic cut sites, put in the same orientation as described in section **2.1**, once again to reduce the incidence of inverse integration by reconstituting the Cas9 cut sites (Fig. 7). The two gRNAs were included with the HITI donor in one plasmid, allowing for a two-plasmid system wherein there was the HITI donor plasmid with two gRNAs and a plasmid containing the SaCas9 expression cassette.

Advantages of this gene editing approach over current therapies include the restoration of a full-length dystrophin protein which is an attractive concept due to the known long-term consequences of a truncated isoform of dystrophin on the functional outcomes of those affected by DMD. Since this approach corrects dystrophin at the genomic level, there is potential for this therapy to be a one-time administration as opposed to therapies that require lifelong dosing to

have the desired effect. Finally, the advantage of this therapy over other therapies that aim at genomic correction is the range of patient cohort that would benefit. Rather than being an approach targeted at a specific mutation, this method would effectively correct any mutation within the exon 41-55 locus of the *DMD* gene.

There are also some advantages this system has that would potentially enhance the therapeutic effect if this strategy gets adapted to humans, the first being the physiology of muscle cells. As muscle progenitor cells mature into mature muscle cells, they fuse, creating myofibers which are multinucleated. This, in turn, lends itself to a gene editing approach, as a small proportion of edited nuclei could have a larger-than-predicted effect. Another advantage is the planned delivery method is the adeno-associated virus serotype 9 (AAV9) which has a tropism for muscle cells.

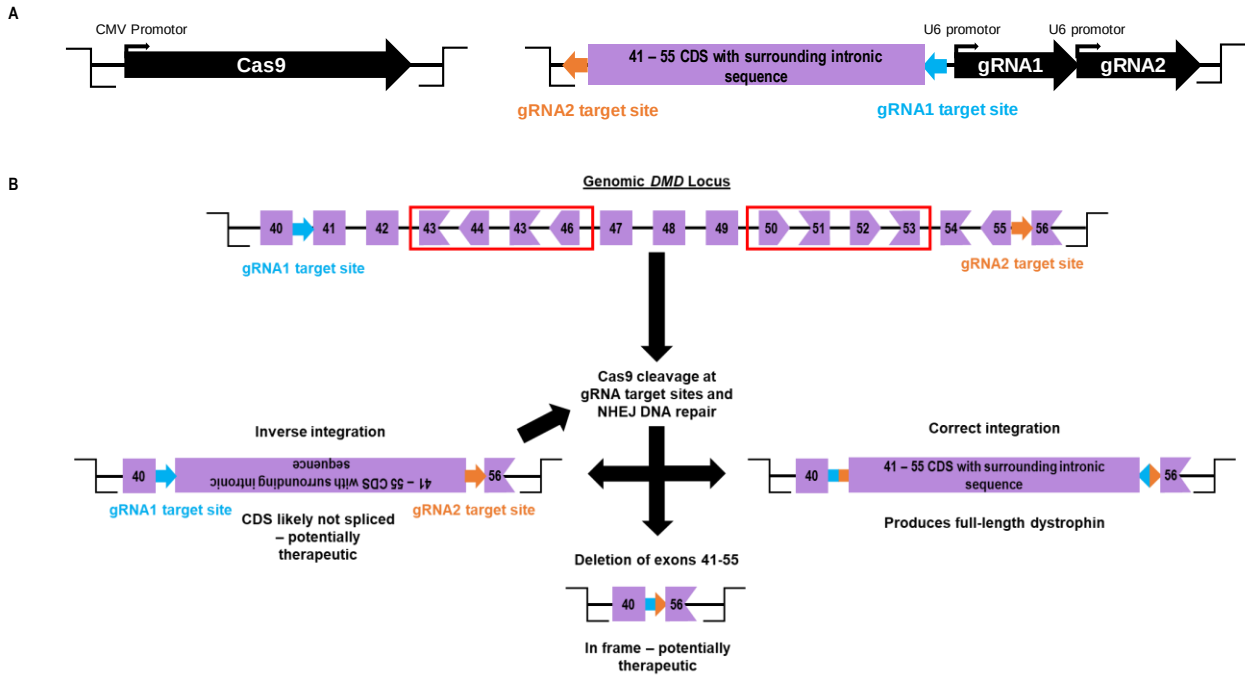

Fig. 7 (**A**) Depiction of the plasmids used for the exon 41-55 HITI replacement strategy with (**B**) a schematic showing the editing outcomes affiliated with this HITI strategy. Red boxes show mutational hotspots within the *DMD* gene.

3.2 Materials and Methods

3.2.1 Molecular Cloning

The plasmids for these experiments were constructed through a variety of cloning methods including inverse PCR, wherein primers are used to amplify an entire plasmid except a portion that is to be deleted. These linearized DNA fragments were then used with the In-Fusion

cloning kit per the manufacturer's recommendations to create new plasmids.

3.2.2 Cell culture and treatments

Culturing of human embryonic kidney 293 (HEK293) cells was accomplished using similar methods as those used in section *2.2.2*. Transfections were also accomplished using Lipofectamine LTX as described in section *2.2.2*, however, the plasmids used were different for these experiments. The transfections in this chapter also used a dual-plasmid transfection system as opposed to the triple plasmid transfection system utilized in the previous chapter.

3.2.3 Fluorescence microscopy

Fluorescence microscopy was accomplished by imaging transfected HEK293 cells at room temperature on a Nikon Ti2-E inverted widefield system with a Hamamatsu Orca Flash 4.0 camera. The dimensions of analyzed images were 1022 X 1024 pixels and were scaled such that there were 1.63 microns/pixel. The fluorescence images were quantified using a custom analysis using the NIS Elements General Analysis 3 module. Cells were identified through automated detection of bright spots of any non-negligible signal intensity after background correction. The mean intensities of red and green signal were then measured and recorded for each bright spot. A threshold based on Otsu methodology was used to differentiate high and low signal in each channel, and spots were counted according to their expression category for each of the two fluorophores. Percent double-positive was calculated by dividing the number of cells with both green and red signal by the total number of cells with both red and green signal, green signal alone, and red signal alone and then multiplying the fraction by 100%.

3.2.4 Genomic DNA PCR analysis

PCR of extracted HEK293 genomic DNA was accomplished using similar methods as those described in section *2.2.3.* Key differences include the primers used, and the cycling conditions. For the experiments conducted in section *3.3.3*, knock-in specific primers were used such that a forward primer was utilized upstream of the genomic Cas9 cut site and a reverse primer was utilized within the HITI knock in. The Tm's were once again calculated based on NEB's Q5 DNA polymerase 2X master-mix online suggestions and the extension time was calculated based on the amplicon length, once again utilizing 30 seconds for every kilobase that was to be amplified.

3.3 Results and Discussion

3.3.1 Identification of active gRNAs

Guide RNAs (gRNAs) targeting upstream of exon 41 (JHI40 series) and downstream of exon 55 (JHI55 series) were designed as described in *2.3.1*. JHI40 series gRNAs targeted within intron 40 as close to exon 40 as possible to include a larger patient cohort (including those with mutations within intron 40). For the JHI55A series gRNAs, an alternative *DMD* gene promoter exists near the 3' end of intron 55 and drives expression of an important dystrophin isoform (Dp116) for Schwann cells.[36] To avoid removing this alternative promoter, JHI55A gRNAs were designed at the 5' end of intron 55, near exon 55. These gRNAs were cloned into plasmids containing an SaCas9 expression cassette driven by the CMVP with the gRNA driven by a U6 promoter as with the experiments described in **Chapter 2**.

The plasmids described above were transfected into HEK293 cells to test the editing capacity of the gRNAs. It was revealed that there were five gRNAs capable of editing from the JHI55A series and three gRNAs were capable of editing from the JHI40 series (Fig. 8). The lead candidates chosen for HITI editing were JHI40-008 and JHI55A-004 (Fig. 8).

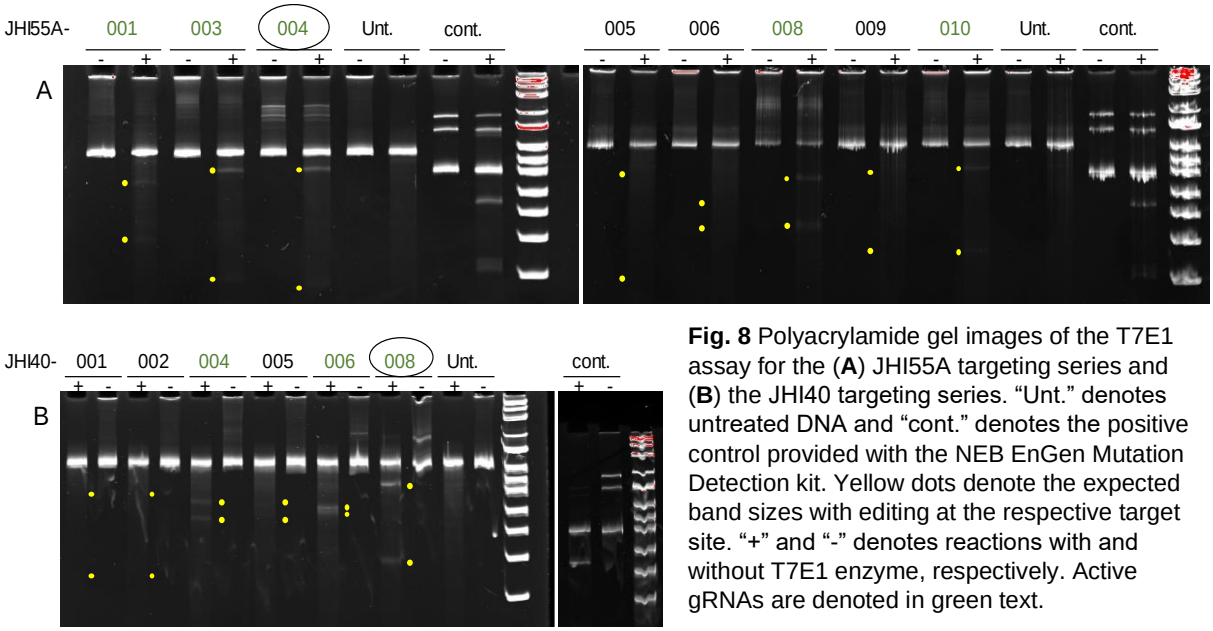

Fig. 8 Polyacrylamide gel images of the T7E1 assay for the (**A**) JHI55A targeting series and (**B**) the JHI40 targeting series. "Unt." denotes untreated DNA and "cont." denotes the positive control provided with the NEB EnGen Mutation Detection kit. Yellow dots denote the expected band sizes with editing at the respective target site. "+" and "-" denotes reactions with and without T7E1 enzyme, respectively. Active gRNAs are denoted in green text.

3.3.2 Co-delivery efficiency

Once the lead gRNAs were chosen, they were cloned into the HITI donor plasmid along with the appropriate Cas9 cut sites on both sides of the donor, while the SaCas9 expression cassette was alone in a separate plasmid. To optimize co-transfection efficiency, the HITI donor

plasmid with gRNAs was tagged with a red fluorescence protein (RFP) and the SaCas9 plasmid was tagged with a green fluorescence protein (GFP) and these plasmids were used to co-transfect HEK293 cells using variable amounts of each plasmid at a 1:1 ratio (0.5 µg, 1.0 µg or 2.0 µg of each plasmid). The cells were imaged using fluorescence microscopy to measure efficiency of co-transfection by the co-localization of RFP and GFP and viability by the estimated percent cell confluency (Fig. 9). Results indicated that the 1.0 µg treated cells had the best co-transfection efficiency with 31.40% double-positive cells, 2.0 µg treated cells had lower

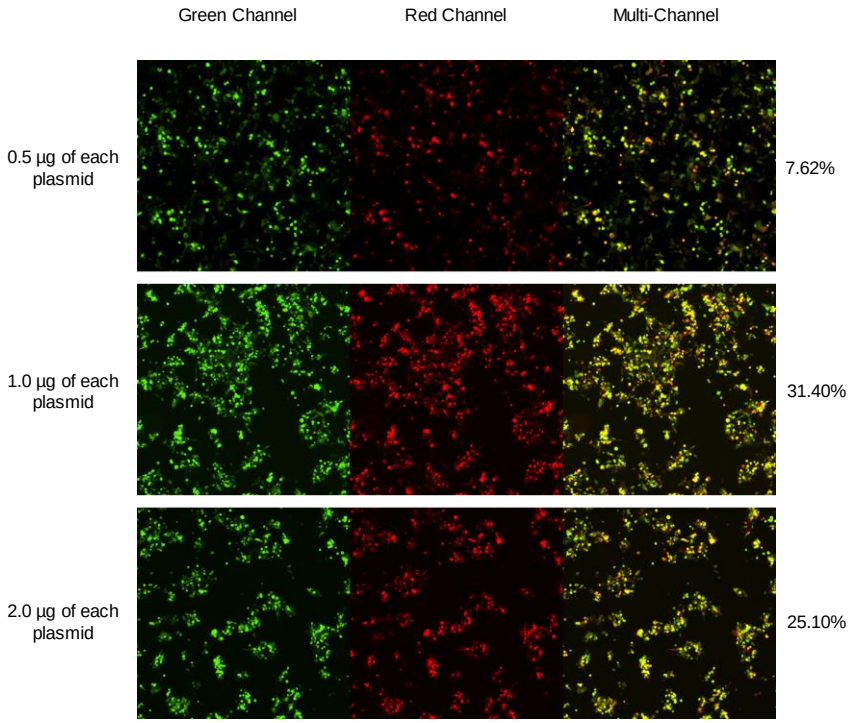

Fig. 9 Fluorescence microscopy images showing co-transfections of the two HITI plasmids with the percentage of double-positive cells compared to the total amount of counted cells.

efficiency at 25.10% double-positive cells, and the 0.5 µg cells had the lowest efficiency with only 7.62% double-positive cells (Fig. 9).

3.3.3 Detection of HITI knock in with the dual-plasmid system

I next wanted to determine whether the dual-plasmid system resulted in proper HITI knock in. Using genomic DNA extracted from the 1.0 µg treated cells described in *3.3.2*, PCR was performed using knock-in specific primers. The results showed successful integration of the HITI donor in the CRISPR and Donor treated cells (Fig. 10A). The CRISPR only treatment and Donor only treatment did not have the expected knock-in band, once again showing that both components are necessary for HITI mediated knock in (Fig. 10A). The untreated cells also did not show the expected knock in band, confirming the lack of knock-in band within the

endogenous genome (Fig. 10A). The knock-in band was sequenced, and this revealed seamless integration (Fig. 10B).

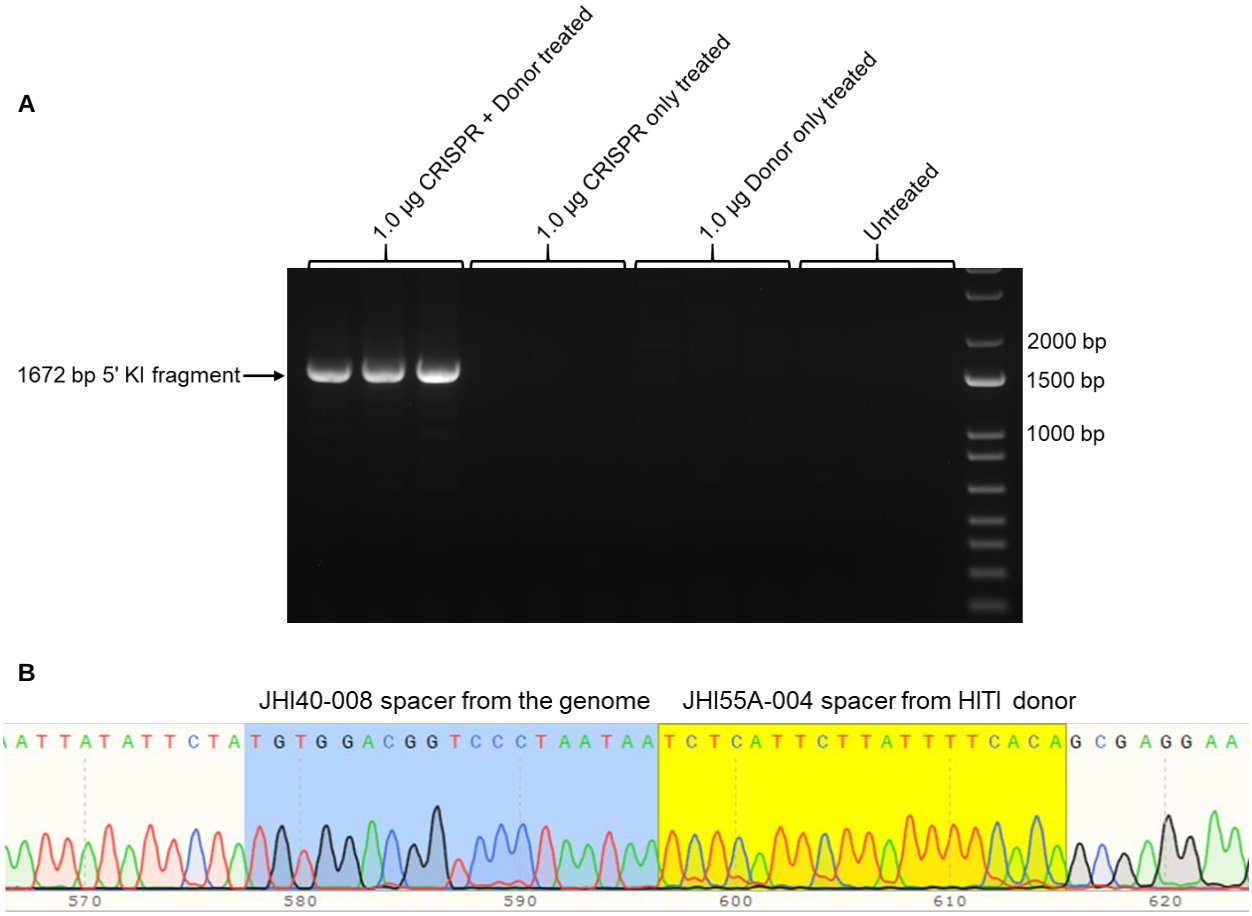

Fig. 10 (**A**) Gel image showing the successful knock in of the HITI donor at the 41-55 site using a 5' knock-in specific primer set. (**B**) Sanger sequencing chromatogram depicting seamless integration of the HITI donor sequence based on the highlighted junction.

3.3.5 Conclusions

These experiments began by designing and confirming gRNAs that targeted an excision of exons 41 through 55, which were subsequently used in experiments for the replacement of those exons using HITI gene editing. The amount of DNA to be transfected was subsequently optimized in HEK293 cells with the new dual-plasmid HITI system to be 1.0 µg. The resulting genomic DNA was utilized in a PCR which showed the successful integration of the HITI donor into the genome. These experiments lay the groundwork for the development of a therapy that has the potential to restore full-length dystrophin while simultaneously reaching a diverse array of DMD-causing mutations.

Chapter 4: Future Directions

4.1 Quantitative Measure of Editing

The immediate next step for this project is to get a robust quantitative measure of the editing efficiency in dual-plasmid transfected cells. To this end I would suggest the use of quantitative PCR (qPCR) with TaqMan technology which would enable a robust quantification of the different editing outcomes and the total amount of *DMD* loci. Different primer/probe pairs would be generated for each editing outcome and the total *DMD* loci, and with these numbers, the editing efficiency could be estimated.

There are also other methods of quantification that could be useful such as digital dropwise PCR (ddPCR) which works similarly to a qPCR but makes use of dividing the PCR reaction into small oil droplets and then quantitating based on the presence of fluorescence or not. This quantitation method, though requiring a more involved workflow, has advantages in that it gives an absolute quantification instead of a quantification based on a standard curve, and also has the ability to detect very low abundance loci in case that is needed.

4.2 Use of cells Derived from DMD Patients

The Flanigan lab has an extensive cell bank containing modified human fibroblast cells derived from DMD patients. These fibroblast cells have been engineered with lentivirus for tetracyline-inducible MyoD expression cassette, which induced differentiation of the fibroblasts into myoblasts and, subsequently, myotubes. Testing the above-performed HITI experiments in these cells would more closely recapitulate the editing environment that would be present within a DMD patient's cells while simultaneously enabling assays to assess expression of the HITI-edited locus. Reverse transcription PCR (RT-PCR) and western blotting could be used to check for correction of the *DMD* mRNA transcript and restoration of dystrophin protein, respectively, while qPCR could be used to quantify the editing at the genomic and mRNA level. Immunofluorescence staining could be used to check for proper localization of dystrophin within the cells.

4.3 Development of DMD HITI Editing for *In Vivo* Experiments

An important step in translating a therapy from the lab into the clinic is showing it works in animal models. Importantly, there exists an animal model of *DMD* that has the human dystrophin gene with mutations existing in the region I am targeting. The presence of a

humanized mouse model is important for my purposes due to the fact that the intronic regions are not homologous between mouse and human, rendering the gRNAs designed for human cells unable to target Cas9 to the specific places within the genome targeted within humans. With the humanized version of the *DMD* gene, this once again makes it possible to target the region of interest with the designed gRNAs.

To prepare for *in vivo* experiments, an appropriate vector must be chosen, and preparations must be made to use the vector decided upon. The vector we have preliminarily chosen is the viral vector known as AAV9. This vector was chosen on account of its efficiency and its tropism for muscle, allowing for more robust delivery to the targeted cells than other vectors. To prepare for the use of AAV, the gene fragments from the Cas9 plasmid and the HITI donor plasmid would need to be cloned into a vector containing the inverted terminal repeat (ITR) regions that are endogenous to the AAV genome, and are imperative for proper packaging of the genome into the AAV capsid. After this cloning, the newly made AAV plasmids would be sent to the virus core for AAV production.

4.4 *In Vivo* Experiments for DMD HITI Editing

After the production of AAV is complete, experiments testing the *in vivo* efficacy would be warranted. To begin, intramuscular injection of the dual-AAV vectors would be done on the humanized DMD mouse model to check for initial efficacy. This would be accomplished by a combination of RT-PCR, western blotting, and immunofluorescence to check for correction of proper mRNA expression, dystrophin expression, and proper dystrophin localization respectively, in the muscles that were injected. The dual-AAV vectors would then be systemically injected to check the universal efficacy with the same experiments being run as above, except with multiple different muscle groups including the heart and diaphragm, two of the most important targets for dystrophin restoration.

Concluding Remarks

This study established the ability of CRISPR/Cas9 used with the HITI methodology to be applied to large replacements of genomic DNA, which has previously not been explored. By establishing this precedent for such large replacements, the groundwork is set for a future DMD therapy that has the potential to restore full-length dystrophin to a vast cohort of patients with diverse mutations. This methodology also boasts the ability to provide gene correction to those patients that previously had no hope for this type of therapy. The combined positive effects of this therapy are much needed in the field, as there is not a single therapy that has the potential to reach as many patients as this one does. Not only that, but this also expands the scope of the gene editing toolbelt as well, showing it is possible to replace large segments of genomic DNA. This methodology could potentially be used to treat other complex genetic diseases that were previously unable to be treated effectively.

References

1 Flanigan, K. M. Duchenne and Becker muscular dystrophies. *Neurol Clin* **32**, 671-688, viii, doi:10.1016/j.ncl.2014.05.002 (2014).

2 Gao, Q. Q. & McNally, E. M. The Dystrophin Complex: Structure, Function, and Implications for Therapy. *Compr Physiol* **5**, 1223-1239, doi:10.1002/cphy.c140048 (2015).

3 Bladen, C. L., Salgado, D., Monges, S., Foncuberta, M. E., Kekou, K., Kosma, K., Dawkins, H., Lamont, L., Roy, A. J., Chamova, T., Guergueltcheva, V., Chan, S., Korngut, L., Campbell, C., Dai, Y., Wang, J., Barisic, N., Brabec, P., Lahdetie, J., Walter, M. C., Schreiber-Katz, O., Karcagi, V., Garami, M., Viswanathan, V., Bayat, F., Buccella, F., Kimura, E., Koeks, Z., van den Bergen, J. C., Rodrigues, M., Roxburgh, R., Lusakowska, A., Kostera-Pruszczyk, A., Zimowski, J., Santos, R., Neagu, E., Artemieva, S., Rasic, V. M., Vojinovic, D., Posada, M., Bloetzer, C., Jeannet, P. Y., Joncourt, F., Diaz-Manera, J., Gallardo, E., Karaduman, A. A., Topaloglu, H., El Sherif, R., Stringer, A., Shatillo, A. V., Martin, A. S., Peay, H. L., Bellgard, M. I., Kirschner, J., Flanigan, K. M., Straub, V., Bushby, K., Verschuuren, J., Aartsma-Rus, A., Beroud, C. & Lochmuller, H. The TREAT-NMD DMD Global Database: analysis of more than 7,000 Duchenne muscular dystrophy mutations. *Hum Mutat* **36**, 395-402, doi:10.1002/humu.22758 (2015).

4 Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. & Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816-821, doi:10.1126/science.1225829 (2012).

5 Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. & Zhang, F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308, doi:10.1038/nprot.2013.143 (2013).

6 Zhang, Y., Long, C., Bassel-Duby, R. & Olson, E. N. Myoediting: Toward Prevention of Muscular Dystrophy by Therapeutic Genome Editing. *Physiol Rev* **98**, 1205-1240, doi:10.1152/physrev.00046.2017 (2018).

7 Ran, F. A., Cong, L., Yan, W. X., Scott, D. A., Gootenberg, J. S., Kriz, A. J., Zetsche, B., Shalem, O., Wu, X., Makarova, K. S., Koonin, E. V., Sharp, P. A. & Zhang, F. In vivo genome editing using Staphylococcus aureus Cas9. *Nature* **520**, 186-191, doi:10.1038/nature14299 (2015).

8 Grieger, J. C. & Samulski, R. J. Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps. *J Virol* **79**, 9933-9944, doi:10.1128/JVI.79.15.9933-9944.2005 (2005).

9 McDonald, C. M., Henricson, E. K., Abresch, R. T., Duong, T., Joyce, N. C., Hu, F., Clemens, P. R., Hoffman, E. P., Cnaan, A., Gordish-Dressman, H. & Investigators, C. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. *Lancet* **391**, 451-461, doi:10.1016/S0140-6736(17)32160-8 (2018).

10 Barber, B. J., Andrews, J. G., Lu, Z., West, N. A., Meaney, F. J., Price, E. T., Gray, A., Sheehan, D. W., Pandya, S., Yang, M. & Cunniff, C. Oral corticosteroids and onset of cardiomyopathy in Duchenne muscular dystrophy. *J Pediatr* **163**, 1080-1084 e1081, doi:10.1016/j.jpeds.2013.05.060 (2013).

11 McNally, E. M., Kaltman, J. R., Benson, D. W., Canter, C. E., Cripe, L. H., Duan, D., Finder, J. D., Groh, W. J., Hoffman, E. P., Judge, D. P., Kertesz, N., Kinnett, K., Kirsch, R., Metzger, J. M., Pearson, G. D., Rafael-Fortney, J. A., Raman, S. V., Spurney, C. F., Targum, S. L., Wagner, K. R., Markham, L. W., Working Group of the National Heart, L., Blood, I. & Parent Project Muscular, D. Contemporary cardiac issues in Duchenne muscular dystrophy. Working Group of the National Heart, Lung, and Blood Institute in collaboration with Parent Project Muscular Dystrophy. *Circulation* **131**, 1590-1598, doi:10.1161/CIRCULATIONAHA.114.015151 (2015).

12 Khan, S. M. & Bennett, J. P., Jr. Development of mitochondrial gene replacement therapy. *J Bioenerg Biomembr* **36**, 387-393, doi:10.1023/B:JOBB.0000041773.20072.9e (2004).
13 Pawlyk, B. S., Smith, A. J., Buch, P. K., Adamian, M., Hong, D. H., Sandberg, M. A., Ali, R. R. & Li, T. Gene replacement therapy rescues photoreceptor degeneration in a murine model of Leber congenital amaurosis lacking RPGRIP. *Invest Ophthalmol Vis Sci* **46**, 3039-3045, doi:10.1167/iovs.05-0371 (2005).
14 Bachrach, E., Li, S., Perez, A. L., Schienda, J., Liadaki, K., Volinski, J., Flint, A., Chamberlain, J. & Kunkel, L. M. Systemic delivery of human microdystrophin to regenerating mouse dystrophic muscle by muscle progenitor cells. *Proc Natl Acad Sci U S A* **101**, 3581-3586, doi:10.1073/pnas.0400373101 (2004).
15 Le Guiner, C., Servais, L., Montus, M., Larcher, T., Fraysse, B., Moullec, S., Allais, M., Francois, V., Dutilleul, M., Malerba, A., Koo, T., Thibaut, J. L., Matot, B., Devaux, M., Le Duff, J., Deschamps, J. Y., Barthelemy, I., Blot, S., Testault, I., Wahbi, K., Ederhy, S., Martin, S., Veron, P., Georger, C., Athanasopoulos, T., Masurier, C., Mingozzi, F., Carlier, P., Gjata, B., Hogrel, J. Y., Adjali, O., Mavilio, F., Voit, T., Moullier, P. & Dickson, G. Long-term microdystrophin gene therapy is effective in a canine model of Duchenne muscular dystrophy. *Nat Commun* **8**, 16105, doi:10.1038/ncomms16105 (2017).
16 Mendell, J. R., Rodino-Klapac, L. R., Sahenk, Z., Roush, K., Bird, L., Lowes, L. P., Alfano, L., Gomez, A. M., Lewis, S., Kota, J., Malik, V., Shontz, K., Walker, C. M., Flanigan, K. M., Corridore, M., Kean, J. R., Allen, H. D., Shilling, C., Melia, K. R., Sazani, P., Saoud, J. B., Kaye, E. M. & Eteplirsen Study, G. Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann Neurol* **74**, 637-647, doi:10.1002/ana.23982 (2013).
17 Aartsma-Rus, A. & Corey, D. R. The 10th Oligonucleotide Therapy Approved: Golodirsen for Duchenne Muscular Dystrophy. *Nucleic Acid Ther*, doi:10.1089/nat.2020.0845 (2020).
18 Alfano, L. N., Charleston, J. S., Connolly, A. M., Cripe, L., Donoghue, C., Dracker, R., Dworzak, J., Eliopoulos, H., Frank, D. E., Lewis, S., Lucas, K., Lynch, J., Milici, A. J., Flynt, A., Naughton, E., Rodino-Klapac, L. R., Sahenk, Z., Schnell, F. J., Young, G. D., Mendell, J. R. & Lowes, L. P. Long-term treatment with eteplirsen in nonambulatory patients with Duchenne muscular dystrophy. *Medicine (Baltimore)* **98**, e15858, doi:10.1097/MD.0000000000015858 (2019).
19 Amoasii, L., Hildyard, J. C. W., Li, H., Sanchez-Ortiz, E., Mireault, A., Caballero, D., Harron, R., Stathopoulou, T. R., Massey, C., Shelton, J. M., Bassel-Duby, R., Piercy, R. J. & Olson, E. N. Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. *Science* **362**, 86-91, doi:10.1126/science.aau1549 (2018).
20 Li, H. L., Fujimoto, N., Sasakawa, N., Shirai, S., Ohkame, T., Sakuma, T., Tanaka, M., Amano, N., Watanabe, A., Sakurai, H., Yamamoto, T., Yamanaka, S. & Hotta, A. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Reports* **4**, 143-154, doi:10.1016/j.stemcr.2014.10.013 (2015).
21 Lim, K. R. Q., Yoon, C. & Yokota, T. Applications of CRISPR/Cas9 for the Treatment of Duchenne Muscular Dystrophy. *J Pers Med* **8**, doi:10.3390/jpm8040038 (2018).
22 Jasin, M. & Rothstein, R. Repair of strand breaks by homologous recombination. *Cold Spring Harb Perspect Biol* **5**, a012740, doi:10.1101/cshperspect.a012740 (2013).
23 Bengtsson, N. E., Hall, J. K., Odom, G. L., Phelps, M. P., Andrus, C. R., Hawkins, R. D., Hauschka, S. D., Chamberlain, J. R. & Chamberlain, J. S. Muscle-specific CRISPR/Cas9 dystrophin gene editing ameliorates pathophysiology in a mouse model for Duchenne muscular dystrophy. *Nat Commun* **8**, 14454, doi:10.1038/ncomms14454 (2017).
24 Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262-1278, doi:10.1016/j.cell.2014.05.010 (2014).

25 Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., Wu, J., Zhu, J., Kim, E. J., Hatanaka, F., Yamamoto, M., Araoka, T., Li, Z., Kurita, M., Hishida, T., Li, M., Aizawa, E., Guo, S., Chen, S., Goebl, A., Soligalla, R. D., Qu, J., Jiang, T., Fu, X., Jafari, M., Esteban, C. R., Berggren, W. T., Lajara, J., Nunez-Delicado, E., Guillen, P., Campistol, J. M., Matsuzaki, F., Liu, G. H., Magistretti, P., Zhang, K., Callaway, E. M., Zhang, K. & Belmonte, J. C. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* **540**, 144-149, doi:10.1038/nature20565 (2016).

26 Zare, K., Shademan, M., Ghahramani Seno, M. M. & Dehghani, H. CRISPR/Cas9 Knockout Strategies to Ablate CCAT1 lncRNA Gene in Cancer Cells. *Biol Proced Online* **20**, 21, doi:10.1186/s12575-018-0086-5 (2018).

27 Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J. & Mateo, J. L. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One* **10**, e0124633, doi:10.1371/journal.pone.0124633 (2015).

28 Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., Karapetyan, K., Katz, K., Liu, C., Maddipatla, Z., Malheiro, A., McDaniel, K., Ovetsky, M., Riley, G., Zhou, G., Holmes, J. B., Kattman, B. L. & Maglott, D. R. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res* **46**, D1062-D1067, doi:10.1093/nar/gkx1153 (2018).

29 Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M. & Sirotkin, K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* **29**, 308-311, doi:10.1093/nar/29.1.308 (2001).

30 Flanigan, K. M., Dunn, D. M., von Niederhausern, A., Soltanzadeh, P., Gappmaier, E., Howard, M. T., Sampson, J. B., Mendell, J. R., Wall, C., King, W. M., Pestronk, A., Florence, J. M., Connolly, A. M., Mathews, K. D., Stephan, C. M., Laubenthal, K. S., Wong, B. L., Morehart, P. J., Meyer, A., Finkel, R. S., Bonnemann, C. G., Medne, L., Day, J. W., Dalton, J. C., Margolis, M. K., Hinton, V. J., United Dystrophinopathy Project, C. & Weiss, R. B. Mutational spectrum of DMD mutations in dystrophinopathy patients: application of modern diagnostic techniques to a large cohort. *Hum Mutat* **30**, 1657-1666, doi:10.1002/humu.21114 (2009).

31 Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R. & Schwede, T. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296-W303, doi:10.1093/nar/gky427 (2018).

32 Bienert, S., Waterhouse, A., de Beer, T. A., Tauriello, G., Studer, G., Bordoli, L. & Schwede, T. The SWISS-MODEL Repository-new features and functionality. *Nucleic Acids Res* **45**, D313-D319, doi:10.1093/nar/gkw1132 (2017).

33 Guex, N., Peitsch, M. C. & Schwede, T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **30 Suppl 1**, S162-173, doi:10.1002/elps.200900140 (2009).

34 Benkert, P., Biasini, M. & Schwede, T. Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27**, 343-350, doi:10.1093/bioinformatics/btq662 (2011).

35 Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L. & Schwede, T. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep* **7**, 10480, doi:10.1038/s41598-017-09654-8 (2017).

36 Matsuo, M., Awano, H., Matsumoto, M., Nagai, M., Kawaguchi, T., Zhang, Z. & Nishio, H. Dystrophin Dp116: A yet to Be Investigated Product of the Duchenne Muscular Dystrophy Gene. *Genes (Basel)* **8**, doi:10.3390/genes8100251 (2017).